

Yale University
EliScholar – A Digital Platform for Scholarly Publishing at Yale

Yale Medicine Thesis Digital Library

School of Medicine

January 2018

Quality Assessment Of Clinical Practice Guidelines In Emergency Medicine

Alyssa B. Zupon

Follow this and additional works at: <https://elischolar.library.yale.edu/ymtdl>

Recommended Citation

Zupon, Alyssa B., "Quality Assessment Of Clinical Practice Guidelines In Emergency Medicine" (2018). *Yale Medicine Thesis Digital Library*. 3469.
<https://elischolar.library.yale.edu/ymtdl/3469>

This Open Access Thesis is brought to you for free and open access by the School of Medicine at EliScholar – A Digital Platform for Scholarly Publishing at Yale. It has been accepted for inclusion in Yale Medicine Thesis Digital Library by an authorized administrator of EliScholar – A Digital Platform for Scholarly Publishing at Yale. For more information, please contact elischolar@yale.edu.

Quality Assessment of Clinical Practice Guidelines in Emergency Medicine

A Thesis Submitted to the Yale University School of Medicine in Partial Fulfillment of
the Requirements for the Degree of Doctor of Medicine

by

Alyssa Brenda Zupon

2018

QUALITY ASSESSMENT OF CLINICAL PRACTICE GUIDELINES IN EMERGENCY MEDICINE
Alyssa Zupon, Ting-Xu Tan, Gina Siddiqui, Katherine Couturier, Matthew James, Craig Rothenberg, Arjun Venkatesh. Department of Emergency Medicine, Yale University School of Medicine, New Haven, CT.

Clinical practice guidelines (CPGs) have been published by the American College of Emergency Physicians (ACEP) since 1990 to advance the delivery of emergency medical care. These guidelines have raised controversy, and recent research shows that they are largely based on lower classes of evidence and expert opinion despite Institute of Medicine recommendations. The rigor of development and overall quality of these guidelines have not yet been assessed. Thus, we performed a systematic review and meta-analysis to evaluate the quality of ACEP Clinical Policies using a recognized, validated appraisal instrument: Appraisal of Guideline for Research & Evaluation (AGREE II).

The AGREE II instrument contains 23 appraisal items (scored on a 1 - 7 scale) in six quality domains and two overall assessments. The domains are Scope and Purpose, Stakeholder Involvement, Rigor of Development, Clarity of Presentation, Applicability, and Editorial Independence. Appraisals were performed independently and in random order by five trained appraisers. Primary outcomes were AGREE II ratings for each item, domain, and Overall Assessment. Domain and Overall Assessment ratings were standardized for analyses. Secondary analyses examined associations between AGREE II ratings and date of publication, strength of underlying evidence, and strength of recommendations. Additional analysis examined relationships between domain and Overall Assessment ratings.

Twenty guidelines published from October 2008 to November 2017 were included. Of the six domains, Scope and Purpose scored highest and varied least (mean 90%, coefficient of variation (CV) 0.03), while Applicability scored lowest and varied most (mean 35%, CV 0.16). The four remaining domains had mean scores of 53% - 78% and CVs of 0.3 - 0.14. The mean Overall Assessment rating was 69% (CV 0.13) and was not associated with CPG publication date, strength of underlying evidence, or strength of recommendations. Statistically significant relationships were found between Overall Assessment ratings and two domains (Rigor of Development and Clarity of Presentation).

Based on validated criteria, ACEP Clinical Policies have identifiable areas of strength and weakness. The overall CPG quality did not improve over time and is not explained by the quality of underlying evidence. ACEP Clinical Policies can be improved by including patient representation in the guideline development process and addressing factors that influence the application of these guidelines in clinical practice.

Acknowledgements

I extend my sincerest thanks to my mentor, Dr. Arjun Venkatesh. I have learned so much during this project not only from your teaching but also from your incredible efficiency, creativity, dedication, and care. I could not have asked for a better role model and guide through this process. Your encouragement, thorough and prompt feedback, and tireless enthusiasm for teaching have made a tremendous difference for me, and I am truly inspired to one day make that same difference for future students.

Thank you, Craig, for all of your statistical expertise and insights, and also many thanks to the emergency medicine residents Ting, Gina, Kate, and Matt. This project would be nothing without you all.

Finally, thank you, Mom and Dad, for your inexhaustible and overflowing support in this project and in every other endeavor I can think of. You have been the foundation underlying my every accomplishment and the fuel feeding my every dream.

Conflict of interest: none

Grant support: Office of Student Research, Yale School of Medicine

Department support: Department of Emergency Medicine, Yale New Haven Hospital

Table of Contents

ABSTRACT	i
ACKNOWLEDGEMENTS	ii
TABLE OF CONTENTS	iii
INTRODUCTION	1
Background	1
Literature Review	1
STATEMENT OF PURPOSE	5
METHODS	6
Overview of thesis responsibilities	6
Study design	6
Selection of clinical practice guidelines.....	6
Data abstraction	7
Outcomes	18
Analysis	18
RESULTS	21
Data set description.....	21
Primary analysis.....	21
Secondary analysis.....	22
DISCUSSION	24
Study summary	24
Discussion of primary analysis	24
Discussion of secondary analysis	30
Implications for AGREE II instrument developers and users	32
Implications for ACEP Clinical Policy development and use	33
LIMITATIONS	34
CONCLUSIONS	35
REFERENCES	36
FIGURES & TABLES	39
APPENDIX	54

Introduction

Background

The explosion of medical literature in recent decades has provided increased access to evidence that can improve health care delivery, yet the volume of available resources can be overwhelming, the quality of the literature uncertain, and the consistency of information incongruous (1-3). Particularly in an era of digitized and electronic resources and libraries, it is possible to quickly disseminate vast quantities of information without standardized organization or curation of that content, thereby limiting translation into practice. Indeed, the Institute of Medicine's (IOM) 2001 report "*Crossing the Quality Chasm: A New Health System for the 21st Century*" identified significant gaps between best-available evidence and medical practice (4).

In an effort to close these gaps, clinical practice guidelines (CPGs) were developed to provide synthesized and critically appraised evidence to enhance clinician and patient decision-making. CPGs aim to translate the complexity of scientific research findings into clinical recommendations and advance the quality of health care delivery through acceleration of knowledge translation, promotion of cost-effective practices, and reduction of practice variation (5). Further, CPGs can influence the development of educational programs, quality measures, and research agendas (6).

Literature review

In the past three decades, there has been a notable increase in the publication and use of CPGs among various medical specialty societies, health care institutions, and governmental bodies (5, 7). In 1989, there were an estimated 700 available guidelines, and, in 2017, there are over 6,400 guidelines from 79 countries listed in the Guidelines

International Network database (1, 8). This increase in the number of guidelines, however, has raised concern about the lack of standardization in the development process and information presentation. Studies have shown that the recommendations provided in various CPGs can be conflicting (9, 10) and the information can be invalid, unreliable, or irrelevant to clinical practice (3, 11). Further, low quality CPGs have the potential to undermine the credibility of more valid guidelines and lead to patient harm if incorrect recommendations were implemented into practice (3, 9, 10). A prior review on the quality of guidelines produced in the 1990s showed that, out of 431 CPGs, 67% did not describe the individuals involved with guideline development, 88% did not report how evidence was identified, and 82% did not indicate the strength of recommendations (3, 11).

Concerns about the low quality and high variability of CPGs motivated the development of several initiatives to formalize methods for guideline appraisal. Most notably is the Appraisal of Guidelines for Research and Evaluation (AGREE) Collaboration, established in 1998 by an international group of researchers from 13 countries (12). In 2003, this group developed the original AGREE instrument, which provided a systematic framework for assessing the methodological rigor, transparency of development, and quality of reporting in CPGs. The AGREE instrument has been translated into numerous languages, is the only appraisal tool that has been validated internationally, and has been formally endorsed by several organizations such as the World Health Organization Advisory Committee on Health Research (3, 12, 13). The instrument underwent revision in 2010, and the new AGREE II instrument has since been cited in over 650 publications (14).

Shortly after the release of the AGREE II instrument, the IOM also published its own report aimed to outline objective, scientifically valid, and consistent approaches to developing practice guidelines (5, 16). This 2011 report, “*Clinical Practice Guidelines We Can Trust*,” provided eight standards of guideline trustworthiness that are similar in content and structure to the AGREE II criteria (5, 12, 16, 17). Both the AGREE II instrument and IOM’s standards of trustworthiness have been widely applied in efforts to advance the quality and appropriate utilization of clinical practice guidelines (17- 27).

Among the many medical specialties promulgating CPGs is emergency medicine. Given the broad nature of emergency medical care, policies outlining standardized approaches can have substantial effect on patient outcomes by promoting evidence-based acute diagnostics and treatments, reducing practice variation, and limiting unnecessary costs and interventions. Emergency physicians have also identified CPGs as essential to providing synthesized, curated content and reassurance of legal protection (28, 29). Guidelines specifically for emergency medical care were first published by the American College of Emergency Physicians (ACEP) in 1990. The first guideline focused on non-traumatic chest pain and presented a general approach to undifferentiated disease, focusing on key aspects of the history and physical exam (1). Since then, the ACEP Clinical Policies Committee was formed and a formal methodology was adopted in parallel with the growing popularity and prominence of evidence-based medicine. Since 1998, ACEP has been publishing evidence-based, peer-reviewed guidelines structured to answer specific, clinically relevant questions considered to be of high frequency or high risk in emergency medicine (1, 2).

Despite the positive effect that ACEP Clinical Policies have on the delivery of

emergency care (19), they have also been a target for scrutiny and a source of controversy. ACEP's 2013 Clinical Policy on intravenous tissue plasminogen activator (tPA) use in acute ischemic stroke was met with skepticism by emergency physicians who felt that the recommendations lacked both adequate supporting evidence and consideration of the adverse outcomes of treatment (17, 28). Three out of eight panelists who developed this Clinical Policy disclosed relevant industry relationships, yet seven had affiliations with the Foundation for Education and Research in Neurological Emergencies, which received all of its donations from drug companies, including alteplase manufacturer, Genentech (17). This concern about industry influence resulted in not only revision of the guideline's recommendations but also substantial reevaluation of ACEP's Clinical Policy development process, rating methodology, and management of conflicts of interest (30). This methodological update was applied to the revised version of the Clinical Policy on tPA use, published in September 2015, and all policies published since. Another recent analysis of ACEP's Clinical Policies revealed that the majority of clinical recommendations are based on lower classes of evidence or expert opinion as opposed to higher classes of evidence such as controlled clinical trials (6). To date, there has been no formal appraisal of ACEP's Clinical Policies to evaluate their quality based on validated criteria.

Statement of Purpose

Accordingly, we sought to assess the methodological rigor, transparency of development, and overall quality of current ACEP Clinical Policies using the AGREE II appraisal instrument. Secondly, we sought to examine whether AGREE II ratings reflected improvement in Clinical Policies' quality over time and, specifically, after ACEP methodological updates in 2015. We also evaluated whether certain aspects of CPG quality were related to the CPG's overall quality assessment. Finally, we examined whether there were associations between AGREE II ratings and the strength of underlying evidence or recommendations in these CPGs.

Methods

Overview of thesis responsibilities

AZ and AV established the purpose and scope of this thesis project. AZ developed the draft data collection tool and study protocol with supervision by AV. AZ utilized the AGREE II instrument to assess all ACEP Clinical Policies in this study. To meet recommendations made by AGREE II developers, the project was expanded to include additional review by emergency medicine residents (TT, GS, KC, MJ). AZ, AV, and CR analyzed and interpreted the data. AZ primarily drafted the manuscript with critical review and revision by CR and AV.

Study design

Systematic review of American College of Emergency Physicians Clinical Policies using the AGREE II appraisal instrument (32, 33).

Selection of clinical practice guidelines

We included all American College of Emergency Physician (ACEP) Clinical Policies listed as “current” as of May 24, 2017, from the ACEP Clinical & Practice Management website, <http://www.acep.org/clinicalpolicies/>. During data collection, one Clinical Policy was replaced with a revised version ("*Emergency Department Management of Patients Needing Reperfusion Therapy for an ST-Segment Elevation Acute Myocardial Infarction*") and thus our study included the revised policy and excluded the prior version. Also, after data collection, one Clinical Policy ("*Critical Issues in the Evaluation and Management of Adult Patients Presenting to the Emergency Department with Syncope*") was removed from the “current” list but remained in this study, as it was current at the time of initial guideline selection. Each Clinical Policy is

comprised of a peer-reviewed manuscript with evidence-based recommendations aimed to guide clinical decision-making. All Clinical Policies are authored by ACEP and follow the ACEP Clinical Policy development process, which includes expert review from medical specialists and societies relevant to the policy topic. These outside participants' affiliations are noted in each Clinical Policy, and their involvement does not imply endorsement of the policy. ACEP Clinical Policies are specific to emergency care in the United States, are regularly published and maintained, and are sponsored by ACEP (6). While other professional organizations publish guidelines for, or relevant to, emergency care, few other groups have a regular process or a committee responsible for guideline maintenance. This study did not include clinical practice guidelines either published by other professional organizations in emergency medicine or primarily authored by other organizations and co-signed or endorsed by ACEP.

Data abstraction

The AGREE II instrument

The data were abstracted using the electronic web tool created by the AGREE II developers, available at <http://www.agreetrust.org/>. This instrument consists of 23 key items organized within six quality domains and two additional global assessments. Each item is rated on a Likert scale between 1 (strongly disagree) and 7 (strongly agree). Each domain captures a unique dimension of guideline quality, specifically Scope and Purpose, Stakeholder Involvement, Rigor of Development, Clarity of Presentation, Applicability, and Editorial Independence.

Scope and Purpose is concerned with the overall aim of the guideline, the specific health questions, and the target population.

Stakeholder Involvement focuses on the extent to which the guideline development group includes individuals from all relevant professional groups, represents the views of its intended users, and clearly defines its target population.

Rigor of Development relates to the process utilized to gather and synthesize the evidence, the methods utilized to formulate the recommendations, and the criteria used to update them. Specifically, the items in this domain evaluate whether systematic methods were used to search for evidence, criteria for selecting the evidence are clearly described, the strengths and limitations of the body of evidence are identified, the methods for formulating the recommendations are clear, the risks and benefits have been considered, there is an explicit link between the recommendations and the supporting evidence, the guideline has been externally reviewed by experts prior to its publication, and a procedure for updating the guideline is provided.

Clarity of Presentation addresses the language, structure, and format of the guideline, specifically whether the recommendations are specific and unambiguous, other options for management are presented, and key recommendations are easily identifiable.

Applicability pertains to factors affecting guideline implementation, strategies to improve uptake, and resource implications of applying the recommendations in practice. These items evaluate whether the guideline describes facilitators and barriers to its application, advice or tools on how the recommendations can be put into practice, potential resource implications of applying the recommendation, and monitoring and/or auditing criteria.

Editorial Independence is concerned with whether the views of the funding body have influenced the content of the guideline and whether competing interests of guideline group members have been recorded and addressed.

After completing assessments for the six domains, the instrument prompts the reviewer for an overall assessment of the guideline (using the same 1 to 7 Likert scale) and a categorical recommendation for use in clinical practice (“yes,” “yes with modification,” or “no”). These two global assessments are based on the reviewer’s overall impression of the guideline and are not calculated from item or domain ratings. A full description of each quality domain, 23 items, and global assessments is provided in Appendix I.

Sample data abstraction

An example of an individual review of a clinical practice guideline using the AGREE II instrument is provided below (completed by reviewer AZ). Each of the 23 items within the six domains is listed with its associated annotation and reason for rating.

Critical Issues in the Diagnosis and Management of the Adult Psychiatric Patient in the Emergency Department Using the AGREE II Instrument

1. Scope and Purpose

1. The overall objective(s) of the guideline is (are) specifically described.

Rating: 6

Page 481, column 1. This clinical policy from ACEP addresses key issues for the diagnosis and management of adult psychiatric patients in the emergency department. A writing subcommittee conducted a systematic review of literature to derive evidence-based recommendations to answer the following clinical questions. The four clinical questions are clearly delineated.

2. The health question(s) covered by the guideline is (are) specifically described.

Rating: 7

The guideline presents four "critical questions" (pages 483-487) in a very clear manner. Further, based on the definition provided and the discussion of each critical question, the context of the critical question relative to overall clinical practice is provided.

3. The population (patients, public, etc.) to whom the guideline is meant to apply is specifically described.

Rating: 7

Page 483, column 1. Inclusion Criteria: This guideline applies to adult patients presenting to the ED with psychiatric symptoms.

Critical Question 4 includes patients with delirium.

Exclusion Criteria: This guideline is not intended to be used for pediatric patients.

Also not intended for patients with delirium in regard to Critical Questions 1, 2, 3.

2. Stakeholder Involvement

4. The guideline development group includes individuals from all relevant professional groups.

Rating: 5

Page 482, column 1. This policy is a product of the ACEP Clinical Policy development process, including expert review. Expert review comments were received from emergency physicians, psychiatrists, members of the American Association for Emergency Psychiatry, the American Association of Community Psychiatrists, and ACEP's Medical Legal committee. Comments were received during a 60-day open-comment period, with notices of the comment period sent in an email to ACEP members, published in EM Today, and posted on the ACEP Web site. The responses were used to further refine and enhance the policy; however, they do not imply endorsement of this clinical policy. Lacks details of

how input from these institutions, members, or experts was used by the guideline development group.

5. The views and preferences of the target population (patients, public, etc.) have been sought.

Rating: 2

No mention of seeking outside opinion (public, patient, etc.) was noted in the guideline. They do reference in the Intro (page 481) how ED are taking the weight of the substantial decline in mental health resource allocation for patients with mental health issues, but no proposal for how to get additional stakeholders involved.

6. The target users of the guideline are clearly defined.

Rating: 7

Page 483, column 1. Scope of Application: This guideline is intended for physicians working in EDs.

3. Rigor of Development

7. Systematic methods were used to search for evidence.

Rating: 6

Page 482, column 1. This Clinical Policy was created after careful review and critical analysis of medical literature and was based on a systematic review of the literature. Searches of MEDLINE, MEDLINE InProcess, Scopus, Web of Science, and the Cochrane Database were performed. All searches were limited to English-language sources, adults, and human studies. Specific key words/phrases, years used in the searches, dates of searches, and study selection were identified under each critical question. In addition, relevant articles from the bibliographies of included studies and more recent articles identified by committee members and reviewers were included.

8. The criteria for selecting the evidence are clearly described.

Rating: 6

Page 482, column 1. As noted in Item 7. Additionally, the evidence was graded (page 482, column 1) and assigned a class (Class I, II, III) based on a predetermined process taking into account design and quality of study – ("Assessment of Classes of Evidence" – Appendix A, Appendix B and evidentiary table) for use in making the recommendation. The guideline describes the translation of evidence class for making the recommendation (Level A, B, C) (page 482, column 2).

9. The strengths and limitations of the body of evidence are clearly described.**Rating: 6**

Strengths and Limitations are mentioned. Examples: Page 482, column 1. All searches were limited to English-language sources, adults and human studies. Page 482, column 1. When literature was not available, consensus of emergency physicians was used. Page 482, column 2. There are certain circumstances in which the recommendation stemming from a body of evidence should not be rated as highly as the individual studies on which they are based. Page 483, column 1. This policy is not intended to be a complete manual on the diagnosis and management of adult psychiatric patients in the ED. Page 483, column 2. This policy is not intended to represent a legal standard of care for emergency physicians. Page 483, column 2. Recommendations not intended to represent the only diagnostic or management options. The guideline defines for the physician those strategies for which medical literature exist to provide support for answers to the critical questions addressed in this policy. Page 483, column 2. For potential benefits and harms of implementing the recommendations, see Appendix D. Page 486, column 2. This clinical policy demonstrates that there is no tool currently available that can be solely used to predict the risk of suicide among patient in the ED.

10. The methods for formulating the recommendations are clearly described.**Rating: 7**

Methodology (page 482). Provides a detailed description of the methodology used, the review process, the citations by years, etc. There was a thorough assessment of the literature and the evidence was graded and assigned a grade (Class I, II, III) based on predetermined process taking into account design and quality of study (Appendix A and B, Evidentiary Table) for use in making the recommendation (Level A, B, C). For the Critical Questions (pages 483-487) there are logical discussions of the problems and literature leading up to a clear and concise "recommendation" for each of the four critical questions.

11. The health benefits, side effects, and risks have been considered in formulating the recommendations.

Rating: 6

For potential benefits and harms of implementation the recommendation, the guideline provides Appendix D. This appendix uses the same format as the critical questions and recommendation, and provides a review of benefit/harm associated with the recommendations. More details could be provided.

12. There is an explicit link between the recommendations and the supporting evidence.

Rating: 7

The four Critical Questions (pages 483-487) addressed in the guideline are clearly delineated and clinical practice recommendations are provided. Research supporting the recommendation is well described. Further, the recommendations are assigned a level (Level A, B, C) based on the strength of evidence (Class I, II, III) and expert opinion (Assessment of Classes of Evidence and Translation of Classes of Evidence to Recommendation Levels) (page 482, column 2). Additionally, inclusion of "future research" for each of the critical questions (page 484, column 1) provides additional insight into the current evidence supporting the clinical recommendation.

13. The guideline has been externally reviewed by experts prior to its publication.

Rating: 5

Page 482, column 1. This policy is a product of the ACEP Clinical Policy development process, including expert review. Expert review comments were received from emergency physicians, psychiatrists, members of the American Association for Emergency Psychiatry, the American Association of Community Psychiatrists, and ACEP's Medical Legal committee. Comments were received during a 60-day open-comment period, with notices of the comment period sent in an email to ACEP members, published in EM Today, and posted on the ACEP Web site. The responses were used to further refine and enhance this policy; however, they do not imply endorsement of this clinical policy. Lacks detailed description of reviewers, outcome and summary of key findings, and description of how/whether information was used in the guideline.

14. A procedure for updating the guideline is provided.

Rating: 5

Page 482, column 1. Clinical policies are scheduled for review and considered for revision every 3 years, however, interim reviews are conducted when technology, methodology or the practice environment changes significantly. No definition/explanation of what "significantly" means.

4. Clarity of Presentation

15. The recommendations are specific and unambiguous.

Rating: 5

Critical Questions, pages 483-487. While the critical questions are clearly delineated, they are occasionally vague, and given the nature of the patients being treated (psychiatric patients) there are provided a range of management strategies. But in general, by grading the level of evidence (Class I, II, III) and using that grading in establishing the recommendation (Level A, B, C) the

guideline provides the ED physician with a level of confidence on how to interpret the recommendation, which is helpful.

16. The different options for management of the condition or health issue are clearly presented.

Rating: 5

Each critical question (pages 483-487) is clearly defined, and research supporting the recommendation delineated. In these discussions, they provide clinical background and alternate choices. For example, Page 484, column 1. Existing literature indicates that routine or ancillary laboratory testing for psychiatric patients has little or no use in ED. It is likely that subsets of patients with higher rates of disease (e.g., elderly, immunosuppressed, new onset psychosis, substance abuse) may benefit from routine laboratory testing. Although urine toxicology screen has no benefit for management or disposition of the patient in the ED, it may be helpful to obtain an objective understanding of the patient's potential substance abuse on transfer to a psychiatric facility.

17. Key recommendations are easily identifiable.

Rating: 7

The critical questions addressed in the guideline (pages 483-487) are clearly delineated and clinical practice recommendations are provided. Further the recommendations are assigned a level (Level A, B, C) based on the strength of evidence (Class I, II, III) and expert opinion (page 482).

5. Applicability

18. The guideline describes facilitators and barriers to its application.

Rating: 4

Overall does not identify facilitators and barriers to application, but makes some comments about application. For example, Page 484, column 1: To expedite the care of patients, agreement between the ED and local psychiatric facilities regarding minimal laboratory testing for psychiatric clearance should be

mutually determined. Page 485, column 1: Given that many acutely psychiatric patients may not be able to cooperate with a comprehensive neurological examination, emergency physicians may have a lower threshold to obtain neuroimaging in these patients. Page 486, column 2: This clinical policy review demonstrates that there is no tool currently available that can be solely used to predict the risk of suicide among patients in the ED who have suicidal ideation.

19. The guideline provides advice and/or tools on how the recommendations can be put into practice.

Rating: 2

The guideline does not address this explicitly. However, since the guideline is provided by ACEP to all members and published in Annals of Emergency Medicine, mere dissemination of the guideline promotes its use. The format of the guideline, by clearly identifying critical questions and then providing recommendations, facilitates use of the guideline.

20. The potential resource implications of applying the recommendations have been considered.

Rating: 2

Resource implications for applying the guidelines were not explicitly addressed. However, in the evaluations and eventual recommendations (pages 483-487) it mentions trying to optimize testing and emergency department time and still provide optimal care.

21. The guideline presents monitoring and/or auditing criteria.

Rating: 2

The guideline does not specifically address monitoring/auditing.

6. Editorial Independence

22. The views of the funding body have not influenced the content of the guideline.

Rating: 5

Page 482, column 1. ACEP was the funding source for this ACEP clinical policy. The members of the ACEP policies committee could thus have an influence on this final guideline. This may be minor, but there is an opening for influence.

23. Competing interests of guideline development group members have been recorded and addressed.

Rating: 5

Page 487, column 2. There were no relevant industry relationships disclosed by the subcommittee members for this topic. Relevant industry relationships are those relationships with companies associated with products or services that significantly impact the specific aspect of disease addressed in the critical questions. Would be more complete if all relationships the authors have both financially and academically were reported.

Overall Assessment

Overall Quality of This Guideline: 5/7

Guideline Recommended for Use? Yes.

Group appraisal process

Data abstraction was designed to meet or exceed guidance of the AGREE II instrument developers. The AGREE II developers recommend a minimum of two appraisers and optimally four appraisers for stable estimates of CPG quality. In our study, all twenty ACEP Clinical Policies were reviewed by five appraisers (AZ, TT, GS, KC, and MJ). Prior to data abstraction, each appraiser completed a standardized online training module specifically for use of AGREE II (12, 34), and a group session was

conducted after reviewing the first three guidelines to ensure consistent use of definitions. All appraisals were performed independently between May 2017 and September 2017. Further, each appraiser was assigned a unique and random order to perform the data abstraction to minimize bias due to increased familiarity with the instrument over time.

Data abstraction for underlying evidence and recommendations

Each Clinical Policy contains clinical recommendations based on medical literature to address critical questions faced by emergency physicians. For each recommendation found in a Clinical Policy, we recorded the proportion of recommendations that were Level C (the weakest level of recommendation). Level C recommendations are based on evidence from Design Class III studies or expert consensus. We also recorded the proportion of references within each Clinical Policy that were graded as Design Class III evidence. Studies considered as Class III evidence are case series, case reports, and consensus or review papers.

Outcomes

The primary outcomes were AGREE II ratings for each item, each domain, Overall Assessment, and recommendation for use in clinical practice.

Analysis

Primary analysis

For the primary descriptive analysis, domain and Overall Assessment ratings were standardized as a percentage according to the following formula recommended by AGREE II developers (32):

(Obtained score – Minimum possible score) / (Maximum possible score – Minimum possible score) x 100

Ratings were calculated at the domain level and not at the item level. Using the example assessment above, the domain of Stakeholder Involvement is comprised of three items (items 4, 5, and 6) with item scores of 5, 2, and 7, respectively. The maximum score for this three-item domain is 21 (3 x 7) and the minimum score is 3 (3 x 1). The obtained score from the above example was 14 (5 + 2 + 7). Thus, this domain's standardized score was 61%, calculated by $(14 - 3) / (21 - 3) \times 100$.

The standardized score using this formula was determined for each domain and Overall Assessment for each reviewer. The standardized percentages from all five reviewers were then reported as the mean, standard deviation, coefficient of variation, and range.

Secondary analysis

Secondary analyses were performed to further assess ACEP Clinical Policies. First, we examined whether AGREE II domain and Overall Assessment ratings changed over the ten-year span (2007-2017) that the Clinical Policies were published. We further evaluated AGREE II ratings before and after the ACEP Clinical Policy methodology was updated in September 2015 using t-tests for each domain and Overall Assessment. Second, we evaluated for any relationship between each domain rating and Overall Assessment rating by calculating correlation coefficients between each domain and the Overall Assessment rating. Third, we examined the association between AGREE II ratings and the strengths of the CPGs' underlying evidence and recommendations. We report correlation coefficients between AGREE II Overall Assessment ratings and 1) the

proportion of Class III evidence within an ACEP Clinical Policy and 2) the proportion of Level C recommendations within an ACEP Clinical Policy. These measures have been previously utilized to describe the strength of emergency medicine clinical practice guidelines (6).

In addition, to examine inter-rater reliability between the five appraisers, we calculated intra-class coefficients (ICC) for each of the six quality domains. This approach is consistent with prior studies utilizing the AGREE II instrument. For all analysis, we considered alpha equal to or less than 0.05 to be statistically significant and we accounted for multiple comparisons using the Bonferroni correction where appropriate (35). Data analysis was performed using R statistical software (version 3.4.2).

Results

Data set description

This study included twenty clinical practice guidelines published by ACEP between April 2007 and November 2017. Of all included guidelines, 13 were published prior to methodological updates in September 2015 and seven (35%, ACEP 1- 6 and ACEP 20) were published after the update (Table I).

Results of primary analysis

The mean results from the standardized scores of all five reviewers are provided in Table II. The results for the primary analysis rated by this author (AZ) using the standardized scoring formula can be found in Appendix II.

Of the six AGREE II domains, Scope and Purpose had the highest mean rating and the lowest variability (mean rating 90%, coefficient of variation (CV) of 0.03, range 84 – 96%). Applicability had the lowest mean rating and highest variability (mean rating 35%, CV of 0.16, range 37 - 77%). The four remaining domains, from highest to lowest rating, were Rigor of Development (mean rating 78%, CV of 0.03, range 73 – 83%), Clarity of Presentation (mean rating 75%, CV of 0.07, range 59 – 82%), Editorial Independence (mean rating 68%, CV of 0.13, range 37 – 77%), and Stakeholder Involvement (mean rating 53%, CV of 0.08, range 46 – 63%).

For the Overall Assessment, the mean rating for all twenty CPGs from the five appraisers was 69% with CV of 0.13 and a range from 50% to 83% (Table II). The three Clinical Policies with the highest mean Overall Assessments ratings were “*Neuroimaging and Decisionmaking in Adult Mild Traumatic Brain Injury in the Acute Setting*” (83%), “*Emergency Department Management of Patients Needing Reperfusion Therapy for an*

ST-Segment Elevation Acute Myocardial Infarction” (80%), and *"Clinical Policy for Well-Appearing Infants and Children Younger Than 2 Years of Age Presenting to the Emergency Department With Fever"* (80%). The three Clinical Policies with the lowest mean Overall Assessment ratings were *“Critical Issues in the Evaluation and Management of Emergency Department Patients With Suspected Appendicitis”* (50%), *“Critical Issues in the Evaluation and Management of Adult Patients in the Emergency Department with Asymptomatic Elevated Blood Pressure”* (50%), and *“Clinical Issues in the Prescribing of Opioids for Adult Patients in the Emergency Department”* (53%).

Regarding whether the five appraisers recommended these guidelines for clinical use, the vast majority of responses were “yes” or “yes with modifications” (Table III). There were, however, “no” responses for Clinical Policies on asymptomatic elevated blood pressure (two “no” recommendations), prescribing opioids for adult patients (two “no” recommendations), suspected appendicitis (one “no” recommendation), and acute carbon monoxide poisoning (one “no” recommendation). These four clinical policies also had the four lowest mean Overall Assessment ratings of all Clinical Policies in this study (50%, 53%, 50% and 63%, respectively). There was low inter-rater reliability with ICC values ranging from 0.01 to 0.07.

Results of secondary analysis

There was no significant relationship between Overall Assessment rating and date of CPG publication (Figure 1). Further, there was no significant improvement in the Overall Assessment ratings or in any of the six domains after updates to ACEP guideline development process in September 2015 (Table IV).

Of the six domains, Rigor of Development and Clarity of Presentation had the strongest association with the Overall Assessment rating ($R = 0.70$, $p < 0.001$; $R = 0.77$, $p < 0.0001$, respectively) (Table V). The other four domains' ratings were not strongly associated with Overall Assessment rating (R : 0.01 - 0.51).

The listing of recommendation level, class of evidence, and Overall Assessment rating for each Clinical Policy is shown in Table VIa and VIb. No significant correlation was evident between Overall Assessment and either the proportion of Level C recommendations ($R = 0.06$, $p = 0.96$) (Figure 2a) or proportion of Class III evidence ($R = 0.01$, $p = 0.79$) (Figure 2b).

Discussion

Study summary

Overall, ACEP Clinical Policies rated highly based on the validated AGREE II instrument of CPG quality. There was, however, variability in the quality of ACEP Clinical Policies based on AGREE II domains, with strengths primarily in Scope and Purpose, Rigor of Development, and Clarity of Presentation; weakness in Applicability; and mixed results in Stakeholder Involvement and Editorial Independence. There were no significant improvements in any domain or Overall Assessment ratings over time. The domains of Rigor of Development and Clarity of Presentation were most strongly associated with Overall Assessment ratings. We did not find that AGREE II ratings were sensitive to the CPG's strength of underlying evidence. These findings carry important implications for emergency medicine clinical guideline developers, the broader clinical practice guideline community, and physicians using these guidelines to provide emergency care.

Discussion of primary results

ACEP Clinical Policies showed highest quality in Scope and Purpose. Compared to prior published research, our finding that this domain scored highest (mean rating 90%) and varied least (CV 3.1%) is not surprising. For example, studies utilizing the AGREE II instrument to evaluate CPGs for hepatocellular carcinoma (37), spinal conditions (23), and intracranial aneurysms (36) similarly reported Scope and Purpose as the highest scoring domain. The quality elements of the Scope of Purpose domain (objectives, clinical questions, and target population) are inherently addressed in ACEP Clinical Policies due to ACEP's standardized and formulaic guideline development

process centered on clear clinical questions.

ACEP Clinical Policies also showed strengths in Rigor of Development (mean rating 78%) and Clarity of Presentation (mean rating 75%). The aim of the Rigor of Development domain is to evaluate whether the methods and guideline development process are thorough and transparent, specifically in searching for evidence, assessing the quality of evidence, and formulating recommendations. Prior research has shown that many guidelines in other specialties lack explanation of the quality of underlying evidence or strength of recommendations (36). ACEP Clinical Policies clearly present the level of each recommendation and provide explanation of the CPG's underlying evidence. This is a strength that demonstrates appropriate methodological transparency in ACEP's guideline development process. However, while ACEP Clinical Policies thoroughly describe the process of searching for evidence, they could be improved with a more direct explanation or specific examples of how the evidence is linked to each specific recommendation. For Clarity of Presentation, we found that ACEP Clinical Policies use clear language, are formatted logically, provide specific recommendations, and consider alternative options for management. Overall, high mean ratings in Scope and Purpose, Rigor of Development, and Clarity of Presentation indicate strong technical quality and clear delivery of information in ACEP Clinical Policies.

The lowest scoring domain across ACEP Clinical Policies was Applicability. The quality elements of Applicability include facilitators and barriers to CPG application, resource implications of the recommendations, and monitoring or audit criteria. Other studies have shown similar weakness in the Applicability domain, and thus it is reasonable that this domain had both the lowest score (mean rating 35%) and highest

variability (CV 0.16). For example, Wang et al. (37) evaluated 40 guidelines related to liver cancer and found that most guidelines failed to address tools required for facilitating guideline implementation, resulting in an average Applicability domain rating of 16% using the AGREE II instrument. Low scores in Applicability could reflect a belief that guideline development and guideline implementation are separate activities. Guideline developers may feel that the organizational barriers and cost implications are better discussed among local administrators who can make more individualized decisions based on local settings or institutional priorities. This seems especially true in the practice of emergency medicine, which is greatly affected by practice location, available resources, and patient demographics. Therefore, statements about appropriate implementation of recommendations and resource implications are difficult to make universally, so the low mean rating and high variability for this domain across the 20 CPGs in this study were expected. The Applicability domain ratings for these ACEP Clinical Policies could be improved with clearer monitoring criteria and discussion of both the resource implications and barriers to guideline implementation. Given the resources and time required to develop these guidelines, it is understandable that frequent revisions are infeasible; however, regular updates in a less formal manner, such as an official online forum or mobile application could be utilized to disseminate new influential evidence that becomes available. This approach has been taken by developers of Guideline Central (available at <https://www.guidelinecentral.com/mobile-and-web-apps/>), a digital resource and mobile application that provides official recommendations from various respected medical associations and is frequently updated to reflect the latest content.

There were two domains that received mixed ratings: Stakeholder Involvement (mean rating 53%) and Editorial Independence (mean rating 69%). While ACEP Clinical Policies are developed in collaboration with individuals from multiple relevant professional groups, they do not explicitly mention seeking the views and preferences of the target population (38). This is a common weakness among CPGs in the U.S.: Armstrong and Bloom (38) recently reviewed 101 guideline development organizations and reported that only 8% consistently require patient and public involvement. Given the increasing emphasis on patient partnership in research and policy making, U.S. guideline development organizations should look to examples such as the National Institute for Clinical Excellence in the U.K. for models on including patient representation in guideline development (5, 38). Editorial Independence assesses whether the CPG's development was influenced by funding sources or interests of the authors to determine whether recommendations are based on the best-available evidence alone or affected by those with conflicts of interest. Editorial independence has been a long-standing source of controversy for CPGs in several specialties. In 2004, new cholesterol guidelines greatly expanded the number of people recommended for treatment, and yet the vast majority of guideline authors were found to have relationships with manufacturers of cholesterol lowering drugs (39). Another survey found that most chairs and co-chairs of clinical policy committees had financial conflicts of interest (21). The risks of poor editorial independence are numerous. Guidelines disseminated by respected organizations can be utilized to create institutional protocols, develop quality measures, inform insurance coverage decisions, and influence the selection of medications on drug formularies (17): the quality and reliability of all of these are threatened by biased guidelines. Further,

CPGs can be viewed as reliable authority in malpractice lawsuits, and this may place pressure on physicians to follow guideline recommendations even if they have concern about industry influence over guideline development. Our results for the Editorial Independence domain (mean rating 69%) indicate that ACEP Clinical Policies could improve in this area but are relatively strong in this domain compared to other guidelines (7, 19, 20, 24-27). To improve the strength of Editorial Independence, more explicit and thorough disclosure about both financial and intellectual conflicts of interest among guideline developers should be included. Further, explanation should be provided about how these conflicts of interest were managed and accounted for (i.e., whether certain competing interests excluded members from specific aspects of guideline development).

In addition to the domain level ratings, the mean Overall Assessment rating for all 20 Clinical Policies was quite high at 69%. Of note, however, the Overall Assessment component is a new addition since the original AGREE instrument and warrants consideration of its nuances. Unlike the 23 appraisal items within the six domains that have detailed criteria to guide appraisers, the Overall Assessment does not include specific directions, and is the most subjective component of the AGREE II instrument. This dichotomy of including both formulaic and subjective elements may be confusing for users and those interpreting the instrument's output. AGREE II developers could offer advice about how to arrive at an Overall Assessment rating, perhaps with suggestions for weighing the six domains or by providing other relevant criteria. If, however, AGREE II developers believe it is important to maintain some area for flexibility and subjectivity that transcends specific criteria, it should at least be made clear to those interpreting AGREE II scores that Overall Assessment ratings are subjective and distinct from the

item-defined domain ratings. Overall, these authors believe there is value to this subjective component of the AGREE II instrument, and our results indicate the global strength of the ACEP Clinical Policy development methodology.

Most ACEP Clinical Policies were recommended for clinical use by our five appraisers, which supports our opinion that ACEP Clinical Policies are of high quality. There were, however, four guidelines that some appraisers did not recommend for clinical use. It is noteworthy that the guidelines that received any “no” responses for recommendation for clinical use also had the lowest Overall Assessment ratings, indicating agreement between these two modes of assessment.

Our results extend the findings of two other recent studies that utilized the AGREE II instrument to evaluate individual ACEP Clinical Policies, revealing similar strengths and weaknesses. First, Pak et al. (40) assessed three guidelines on management of hypertension, including one published by ACEP. Compared to our results for that ACEP Clinical Policy, Pak et al. reported a similar average domain rating (67% vs. our result of 65%) and similar rankings for domain scores, with Scope and Purpose and Rigor of Development scoring highest and Applicability scoring lowest. The similar average domain rating and domain ranking suggest good agreement and consistency between two different groups of appraisers evaluating the same ACEP Clinical Policy.

Second, Patel et al. (27) evaluated 24 CPGs related to management of traumatic brain injury, including one ACEP Clinical Policy. They reported mean domain and Overall Assessment ratings similar to those in our study, again suggesting consistency and reliability with the AGREE II instrument between two different appraisal groups. An older study, Tavender et al. (41), utilized the original AGREE instrument in 2010 to

evaluate the same ACEP Clinical Policy on traumatic brain injury. Tavender et al.'s ratings were consistently lower than those in our study, but the order of domain performance, from highest to lowest, was identical to ours, suggesting that Tavender et al.'s appraisers were consistently more strict but identified similar areas of strength and weakness in the Clinical Policy.

Inter-rater reliability

Our results of low intraclass coefficients suggest poor inter-rater reliability. The ICC in this study can be explained by the low variance of ratings in comparison to the total possible variance of a seven-item scale. While a modestly higher ICC may be achieved with a larger number of guidelines or reviewers, the effect of a larger sample size on the ICC is not likely to be substantial as this study already exceeded the recommended number of reviewers by the AGREE II developers. Given the ICC's sensitivity to low score variability between reviewers in each domain, an alternative approach to measuring agreement between reviewers may be more appropriate for examining broad consistency between reviewers (such as whether reviewers consistently rate each domain above or below an average as opposed to a measure of scale).

Discussion of secondary analysis

We conducted several secondary analyses to explore trends and underlying causes of the variation in AGREE II domain and Overall Assessment ratings.

Trends in AGREE II ratings over time

We did not find that domain or Overall Assessment ratings improved over the prior decade or after recent methodological updates. As discussed previously, concern about conflict of interest in the 2013 tPA Clinical Policy generated substantial

controversy and sparked ACEP's methodological updates that became effective in 2015. However, our study using the AGREE II instrument did not find significant improvement in the editorial independence of ACEP's Clinical Policy development process since these updates.

Relationships between domain ratings and overall assessment ratings

To better understand how appraisers may arrive at Overall Assessment ratings, we examined whether particular domain ratings were correlated with Overall Assessment ratings. Our findings that Rigor of Development and Clarity of Presentation were most strongly associated with Overall Assessment do not imply that these two domains are predictive of the Overall Assessment rating, but rather suggest that strong performance on these two domains may be influential on the appraiser's overall impression of the CPG's quality. Similarly, Hoffmann-Esser et al. (42) evaluated various aspects of the AGREE II Overall Assessment ratings in 1453 guidelines and found that Rigor of Development had the strongest correlation with Overall Assessment rating. This likely reflects that clinicians highly value the thoroughness of the literature search and final structure of a guideline in evaluating overall CPG quality.

Strength of underlying evidence and level of recommendations

We found that AGREE II ratings were not sensitive to the proportion of underlying Class III evidence or proportion of Level C recommendations within the CPGs, and this is consistent with the design of the AGREE II instrument. While AGREE II evaluates whether the guideline reports the quality of underlying evidence and presents the recommendations clearly, the instrument does not consider the strength of evidence or recommendations in formulating its quality assessment of the CPG. In other words, the

AGREE II criteria are blind to the quality of evidence supporting the guideline. This is both a strength and weakness of the AGREE II instrument. One benefit is that the AGREE II instrument specifically examines elements performed by, and in the control of, guideline authors, and the instrument is less sensitive to factors beyond authors' control (i.e., amount of available literature or strength of scientific evidence). However, this attribute is a weakness because AGREE II ratings do not necessarily indicate whether the recommendations within a CPG can be followed reliably. It is possible that a CPG can score very high on most criteria of AGREE II and yet have no reliable benefit in clinical practice.

Implications for AGREE II instrument developers and users

Overall, our appraisers found the AGREE II instrument to be well structured, user friendly, thorough, and useful in assessing guideline quality. Notable strengths include its informative online training module, detailed instructions with examples for scoring each item, logical and comprehensive structure of items and domains, and an easy-to-use online interface. A weakness of this instrument for those interested in implementing recommendations into practice is its inability to evaluate the quality of evidence or reliability of the CPG's recommendations. This, however, is advantageous for guideline authors who strictly seek to know if their development process is rigorous and transparent. The current AGREE II reflects methodological processes and not necessarily content, and so high AGREE II scores reflect the quality of reporting more than quality of evidence or content. Including criteria that considers the strength of underlying evidence would create a more comprehensive instrument that could provide physicians with more confidence in implementing CPGs with high AGREE II ratings.

Finally, while the AGREE II instrument has been thoroughly validated and shown to be reliable, evidence is still lacking about whether high AGREE II ratings translate to substantial benefits for guidelines' stakeholders, which include institutions, guideline developers, administrators, physicians, and patients. For example, in the previously discussed Pak et al. study on guidelines for hypertension, ACEP's Clinical Policy had higher AGREE II ratings than the CPG published by the European Society of Hypertension (ESH). Despite the ACEP Clinical Policy's relatively strong performance according to AGREE II domain ratings, the authors of that study did not recommend ACEP's guideline for clinical use, while they did recommend the lower scoring CPG by ESH. Those authors justified this by explaining that the ACEP Clinical Policy did not address management of hypertensive emergency, indicating that AGREE II ratings and clinical usefulness are not necessarily related. Future research should explore how utilization of the AGREE II instrument can affect the implementation of CPGs, knowledge translation, and clinical outcomes.

Implications for ACEP Clinical Policies development and use

Our study has identified specific strengths and weaknesses of ACEP Clinical Policies. Even though addressing all issues of applicability may be infeasible or an inefficient utilization of limited ACEP resources, some additional consideration of resource implications, barriers or facilitators for implementation, and monitoring or audit criteria would improve these Clinical Policies within the current scope of development. Future guidelines should at least inform users of the need to consider applicability issues when implementing guideline recommendations. ACEP Clinical Policies could be also improved with broader stakeholder involvement that includes representation from

patients or target populations and more thorough disclosure about both financial and intellectual conflicts of interest among guideline developers.

Limitations

There are limitations of this project that warrant discussion. While this study has focused on guideline appraisal using the AGREE II instrument, it is worthwhile to compare this instrument with the IOM's standards of guideline trustworthiness in "*Clinical Practice Guidelines We Can Trust*" (5). While some consider the IOM's standards to be more comprehensive than AGREE II (16, 17), the disadvantages to sole utilization of the IOM's standards are that they are not structured into a validated instrument for easy use and they offer only an inflexible all-or-nothing definition of guideline trustworthiness. The AGREE II instrument covers similar content as the IOM's standards in a more user-friendly structure and more realistically manages partial adherence to its criteria and, thus, is best suited for the aims of this study.

Use of the AGREE II instrument also has potential for bias as appraisers become more familiar with the instrument. Ratings performed initially may be affected by a reviewer's unfamiliarity with the instrument or scale. Later, reviewers may improve their consistency, be influenced by their prior ratings, or alter their interpretation of the rating scale. Thus, we aimed to minimize this potential bias by assigning each appraiser a unique randomized order to perform appraisals. In addition, all appraisers were from the same academic center, and this may result in an institutional bias. One analytic limitation of this study approach is the risk of committing type 1 error when performing multiple comparisons; thus, we accounted for this by implementing the Bonferroni correction where appropriate. Based on the ICC calculation, there was low inter-rater reliability in

this study, and this was likely due to a low variance of ratings in comparison to the total possible variance of a seven-item scale. Future work should explore alternative statistical methods for assessing the consistency of responses between reviewers. Finally, our analysis was limited to clinical practice guidelines developed by ACEP and did not include any guidelines published by other specialty societies, in other countries, or in other languages; therefore, many emergency care guidelines housed within the documents and writing of other specialties or organizations are absent from this work.

Conclusions

ACEP Clinical Policies demonstrate strengths and weaknesses based on validated criteria provided by the AGREE II instrument. Guideline quality did not improve over time or after ACEP methodological updates in 2015 and is not related to the quality of underlying evidence. ACEP Clinical Policies can be improved by including patient representation in the guideline development process, increasing editorial independence and transparency, and addressing factors that influence the application of these guidelines in clinical practice.

References

1. Schriger DL, Cantrill SV, and Greene CS. The origins, benefits, harms, and implications of emergency medicine clinical policies. *Ann Emerg Med*. 1993;22(3):597-602.
2. Napoli AM, and Jagoda A. Clinical policies: their history, future, medical legal implications, and growing importance to physicians. *J Emerg Med*. 2007;33(4):425-32.
3. Grilli R, Magrini N, Penna A, Mura G, and Liberati A. Practice guidelines developed by specialty societies: the need for a critical appraisal. *Lancet*. 2000;355(9198):103-6.
4. Institute of M. Crossing the Quality Chasm: A New Health System for the 21st Century. Washington, DC: *National Academies Press*. 2001.
5. Institute of Medicine. Clinical Practice Guidelines We Can Trust. Washington DC: *National Academies Press*. 2011.
6. Venkatesh AK, Savage D, Sandefur B, Bernard KR, Rothenberg C, and Schuur JD. Systematic review of emergency medicine clinical practice guidelines: Implications for research and policy. *PLoS ONE*. 2017;12(6):e0178456.
7. Alonso P, Irfan A, Solà I, Gich I, Delgado-Noguera M, Rigau D, et al. The quality of clinical practice guidelines over the last two decades: a systematic review of guideline appraisal studies. *Qual Saf Health Care*. 2010;19(6):e58.
8. Guideline International Network. <http://www.g-i-n.net/library/international-guidelines-library>. Accessed December 13, 2017.
9. Thomson R, McElroy H, and Sudlow M. Guidelines on anticoagulant treatment in atrial fibrillation in Great Britain: variation in content and implications for treatment. *BMJ*. 1998;316(7130):509-13.
10. Unwin N, Thomson R, O'Byrne AM, Laker M, and Armstrong H. Implications of applying widely accepted cholesterol screening and management guidelines to a British adult population: cross sectional study of cardiovascular disease and risk factors. *BMJ*. 1998;317(7166):1125-30.
11. Varonen H, and Mäkelä M. Practice guidelines in Finland: availability and quality. *Qual Health Care*. 1997;6(2):75-9.
12. Brouwers MC, Kho ME, Browman GP, Burgers JS, Cluzeau F, Feder G, et al. AGREE II: advancing guideline development, reporting and evaluation in health care. *J Clin Epidemiol*. 2010;63(12):1308-11.
13. Guyatt G, and Vandvik PO. Creating clinical practice guidelines: problems and solutions. *Chest*. 2013;144(2):365-7.
14. AGREE Resource Center. <https://www.agreetrust.org/resource-centre/citations-of-core-publications/>. Accessed December 5, 2017.
15. Graham R, Mancher M, Miller Wolman D, Greenfield S, and Steinberg E. Clinical Practice Guidelines We Can Trust. Washington, DC: 2011.
16. Ransohoff DF, Pignone M, and Sox HC. How to decide whether a clinical practice guideline is trustworthy. *JAMA*. 2013;309(2):139-40.
17. Eady EA, Layton AM, Sprakel J, Arents BWM, Fedorowicz Z, and van Zuuren EJ. AGREE II assessments of recent acne treatment guidelines: how well do they

- reveal trustworthiness as defined by the U.S. Institute of Medicine criteria? *Br J Dermatol*. 2017.
18. Bennett WL, Odelola OA, Wilson LM, Bolen S, Selvaraj S, Robinson KA, et al. Evaluation of guideline recommendations on oral medications for type 2 diabetes mellitus: a systematic review. *Ann Intern Med*. 2012;156(1 Pt 1):27-36.
 19. Elesber AA, Decker WW, Smars PA, Hodge DO, Shen WK, and Physicians ACoE. Impact of the application of the American College of Emergency Physicians recommendations for the admission of patients with syncope on a retrospectively studied population presenting to the emergency department. *Am Heart J*. 2005;149(5):826-31.
 20. Zhang X, Zhao K, Bai Z, Yu J, and Bai F. Clinical Practice Guidelines for Hypertension: Evaluation of Quality Using the AGREE II Instrument. *Am J Cardiovasc Drugs*. 2016;16(6):439-51.
 21. Kung J, Miller RR, and Mackowiak PA. Failure of Clinical Practice Guidelines to Meet Institute of Medicine Standards Two More Decades of Little, If Any, Progress. *Archives of Internal Medicine*. 2012;172(21):1628-33.
 22. Lin I, Wiles LK, Waller R, Goucke R, Nagree Y, Gibberd M, et al. Poor overall quality of clinical practice guidelines for musculoskeletal pain: a systematic review. *Br J Sports Med*. 2017.
 23. Yaman ME, Gudeloglu A, Senturk S, Yaman ND, Tolunay T, Ozturk Y, et al. A critical appraisal of the North American Spine Society guidelines with the Appraisal of Guidelines for Research and Evaluation II instrument. *Spine J*. 2015;15(4):777-81.
 24. Chiappini E, Bortone B, Galli L, and de Martino M. Guidelines for the symptomatic management of fever in children: systematic review of the literature and quality appraisal with AGREE II. *BMJ Open*. 2017;7(7):e015404.
 25. Yang C, Zhang Z, Zhang L, Tian J, Yu D, Wang J, et al. Quality assessment of clinical practice guidelines on tic disorders with AGREE II instrument. *Psychiatry Res*. 2018;259:385-91.
 26. Vande Veegaete A, Borra V, De Buck E, and Vandekerckhove P. Guidelines of the International Federation of Red Cross and Red Crescent Societies: an overview and quality appraisal using AGREE II. *BMJ Open*. 2016;6(9):e011744.
 27. Patel A, Vieira MM, Abraham J, Reid N, Tran T, Tomecsek K, et al. Quality of the Development of Traumatic Brain Injury Clinical Practice Guidelines: A Systematic Review. *PLoS One*. 2016;11(9):e0161554.
 28. Mehlman MJ. Medical practice guidelines as malpractice safe harbors: illusion or deceit? *J Law Med Ethics*. 2012;40(2):286-300.
 29. Evidence-Based Medicine and the Law: The Courts Confront Clinical Practice Guidelines. <http://archive.ahrq.gov/research/findings/evidence-based-reports/jhpl/rosoff2.html>. Accessed January 10, 2018.
 30. Radecki R. ACEP Clinical Policy on Intravenous Tissue Plasminogen for Stroke Continues to Evolve. <http://www.acepnow.com/article/acep-clinical-policy-on-intravenous-tissue-plasminogen-for-stroke-continues-to-evolve/>. Accessed December 17, 2017.
 31. Burls A. AGREE II—improving the quality of clinical care. *The Lancet*. 2010;376(9747):1128 - 9.

32. Brouwers M, Kho ME, Browman GP, Burgers JS, Cluzeau F, Feder G, et al. AGREE II: Advancing guideline development, reporting and evaluation in healthcare. *Can Med Assoc J.* 2010.
33. The AGREE II Instrument (Electronic version). <https://www.agreetrust.org/my-agree/>.
34. AGREE II Online Training Tools. <https://www.agreetrust.org/resource-centre/agree-ii-training-tools/>.
35. Armstrong RA. When to use the Bonferroni correction. *Ophthalmic Physiol Opt.* 2014;34(5):502-8.
36. Li Q, Yang Y, Pan Y, Duan L, and Yang H. The quality assessment of clinical practice guidelines for intracranial aneurysms: a systematic appraisal. *Neurosurg Rev.* 2017.
37. Wang Y, Luo Q, Li Y, Wang H, Deng S, Wei S, et al. Quality assessment of clinical practice guidelines on the treatment of hepatocellular carcinoma or metastatic liver cancer. *PLoS One.* 2014;9(8):e103939.
38. Armstrong MJ, Bloom JA. Patient involvement in guidelines is poor five years after institute of medicine standards: review of guideline methodologies. *Res Involv Engagem.* 2017;3:19.
39. Abramson J, Barnard R, Barry H, Bezruchka S, Brody H, and Brown D. Petition to the National Institutes of Health seeking and independent review panel to re-evaluate the national cholesterol education project guidelines. 2004.
40. Pak KJ, Hu T, Fee C, Wang R, Smith M, and Bazzano LA. Acute hypertension: a systematic review and appraisal of guidelines. *Ochsner J.* 2014;14(4):655-63.
41. Tavender EJ, Bosch M, Green S, O'Connor D, Pitt V, Phillips K, et al. Quality and consistency of guidelines for the management of mild traumatic brain injury in the emergency department. *Acad Emerg Med.* 2011;18(8):880-9.
42. Hoffmann-Eßer W, Siering U, Neugebauer EA, Brockhaus AC, Lampert U, and Eikermann M. Guideline appraisal with AGREE II: Systematic review of the current evidence on how users handle the 2 overall assessments. *PLoS One.* 2017;12(3):e0174831.

Figures & Tables

Figure 1. Comparison of Publication Year and Overall Assessment Rating

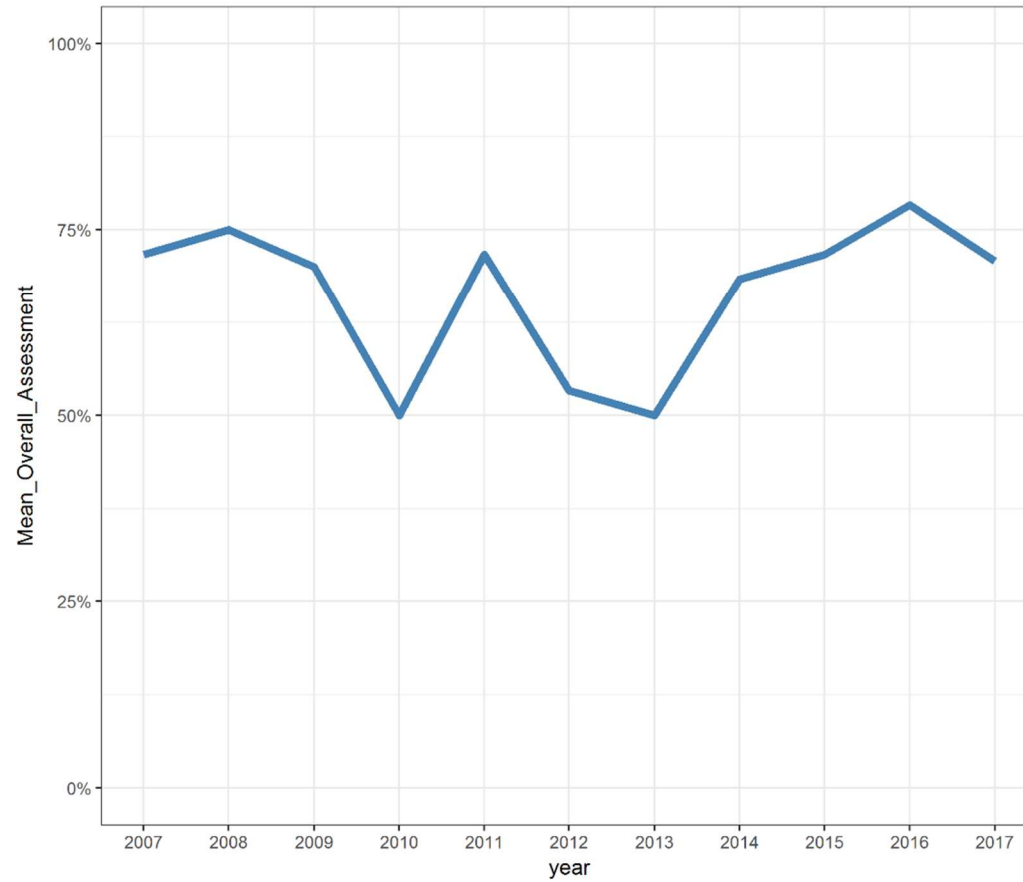
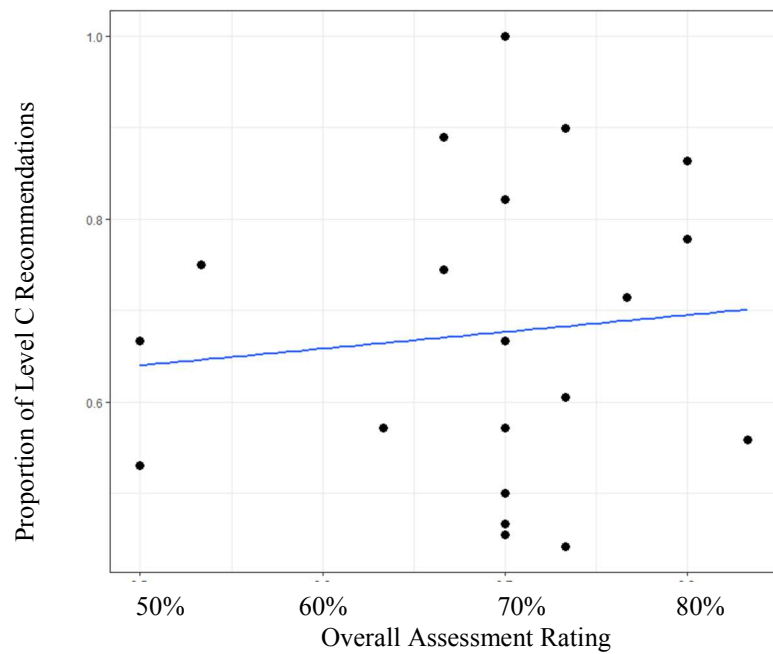


Figure 2. Comparison of Overall Assessment Rating and Either Proportion of Level C Recommendations (2A) or Proportion of Class III Evidence (2B)

A:



B:

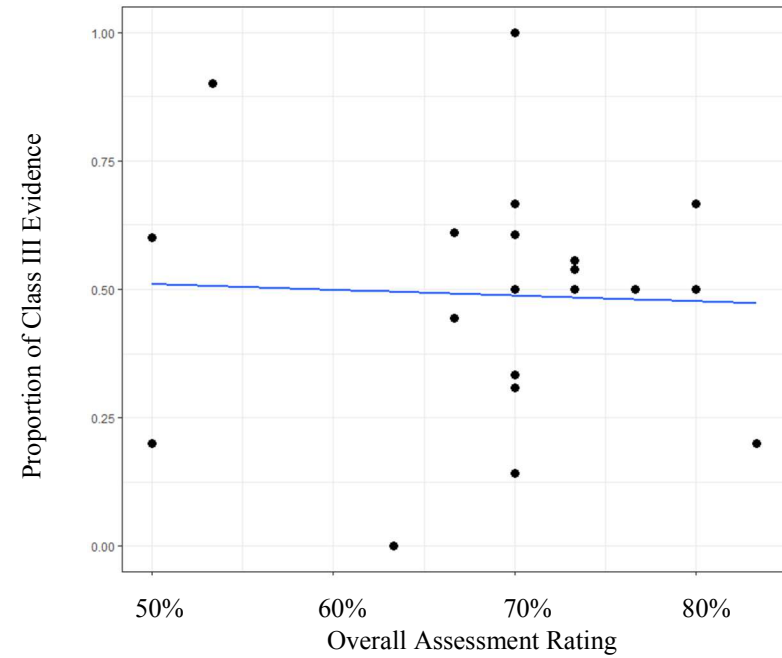


Table I. American College of Emergency Physicians – Current Clinical Policies		
ID	Title	Publication Date
ACEP-1	Critical Issues in the Diagnosis and Management of the Adult Psychiatric Patient in the Emergency Department	Apr 2017
ACEP-2	Critical Issues in the Initial Evaluation and Management of Patients Presenting to the Emergency Department in Early Pregnancy	Feb 2017
ACEP-3	Critical Issues in the Evaluation and Management of Adult Patients Presenting to the Emergency Department with Acute Carbon Monoxide Poisoning	Jan 2017
ACEP-4	Critical Issues in the Evaluation of Adult Patients with Suspected Transient Ischemic Attack in the Emergency Department	Sep 2016
ACEP-5	Clinical Policy for Well-Appearing Infants and Children Younger Than 2 Years of Age Presenting to the Emergency Department With Fever	May 2016
ACEP-6	Use of Intravenous Tissue Plasminogen Activator for the Management of Acute Ischemic Stroke in the Emergency Department	Sep 2015
ACEP-7	Critical Issues in the Evaluation and Management of Adult Patients With Suspected Acute Nontraumatic Thoracic Aortic Dissection	Jan 2015
ACEP-8	Critical Issues in the Evaluation and Management of Adult Patients Presenting to the Emergency Department with Seizures	Apr 2014
ACEP-9	Procedural Sedation and Analgesia in the Emergency Department	Feb 2014
ACEP-10	Critical Issues in the Evaluation and Management of Adult Patients in the Emergency Department with Asymptomatic Elevated Blood Pressure	Jul 2013
ACEP-11	Clinical Issues in the Prescribing of Opioids for Adult Patients in the Emergency Department	Oct 2012
ACEP-12	Critical Issues in the Evaluation and Management of Adult Patients Presenting to the Emergency Department With Suspected Pulmonary Embolism	Jun 2011
ACEP-13	Critical Issues in the Evaluation of Adult Patients Presenting to the Emergency Department With Acute Blunt Abdominal Trauma	Apr 2011
ACEP-14	Critical Issues in the Evaluation and Management of Emergency Department Patients With Suspected Appendicitis	Jan 2010
ACEP-15	Critical Issues in the Management of Adult Patients Presenting to the Emergency Department With Community-Acquired Pneumonia	Nov 2009
ACEP-16	Neuroimaging and Decision making in Adult Mild Traumatic Brain Injury in the Acute Setting	Dec 2008
ACEP-17	Critical Issues in the Evaluation and Management of Adult Patients Presenting to the Emergency Department With Acute Headache	Oct 2008
ACEP-18	Critical Issues in the Evaluation and Management of Adult Patients Presenting to the Emergency Department with Acute Heart Failure Syndromes	May 2007
ACEP-19	Critical Issues in the Evaluation and Management of Adult Patients Presenting to the Emergency Department with Syncope	Apr 2007
ACEP-20	Emergency Department Management of Patients Needing Reperfusion Therapy for an ST-Segment Elevation Acute Myocardial Infarction	Nov 2017

Table II. Assessment of the Agree II Six Quality Domains and Overall Assessment of Clinical Policy									
Guideline Title	ID	Scope and Purpose	Stakeholder Involvement	Rigor of Development	Clarity of Presentation	Applicability	Editorial Independence		Overall Assessment
Critical Issues in the Diagnosis and Management of the Adult Psychiatric Patient in the Emergency Department	ACEP-1	84%	52%	78%	73%	42%	75%		70%
Critical Issues in the Initial Evaluation and Management of Patients Presenting to the Emergency Department in Early Pregnancy	ACEP-2	88%	54%	80%	74%	32%	72%		70%
Critical Issues in the Evaluation and Management of Adult Patients Presenting to the Emergency Department with Acute Carbon Monoxide Poisoning	ACEP-3	86%	51%	76%	78%	32%	77%		63%
Critical Issues in the Evaluation of Adult Patients with Suspected Transient Ischemic Attack in the Emergency Department	ACEP-4	93%	48%	81%	76%	46%	72%		77%
Clinical Policy for Well-Appearing Infants and Children Younger Than 2 Years of Age Presenting to the Emergency Department With Fever	ACEP-5	96%	52%	81%	82%	38%	75%		80%
Use of Intravenous Tissue Plasminogen Activator for the Management of Acute Ischemic Stroke in the Emergency Department	ACEP-6	92%	46%	77%	72%	36%	73%		73%

Table II. Assessment of the Agree II Six Quality Domains and Overall Assessment of Clinical Policy									
Guideline Title	ID	Scope and Purpose	Stakeholder Involvement	Rigor of Development	Clarity of Presentation	Applicability	Editorial Independence		Overall Assessment
Critical Issues in the Evaluation and Management of Adult Patients With Suspected Acute Nontraumatic Thoracic Aortic Dissection	ACEP-7	91%	51%	80%	73%	36%	77%		70%
Critical Issues in the Evaluation and Management of Adult Patients Presenting to the Emergency Department with Seizures	ACEP-8	90%	51%	78%	76%	27%	62%		67%
Procedural Sedation and Analgesia in the Emergency Department	ACEP-9	92%	58%	82%	77%	35%	62%		70%
Critical Issues in the Evaluation and Management of Adult Patients in the Emergency Department with Asymptomatic Elevated Blood Pressure	ACEP-10	90%	49%	74%	69%	33%	77%		50%
Clinical Issues in the Prescribing of Opioids for Adult Patients in the Emergency Department	ACEP-11	90%	63%	75%	72%	36%	67%		53%
Critical Issues in the Evaluation and Management of Adult Patients Presenting to the Emergency Department With Suspected Pulmonary Embolism	ACEP-12	91%	51%	77%	79%	33%	73%		73%
Critical Issues in the Evaluation of Adult Patients Presenting to the Emergency Department With Acute Blunt Abdominal Trauma	ACEP-13	88%	52%	77%	78%	32%	68%		70%

Table II. Assessment of the Agree II Six Quality Domains and Overall Assessment of Clinical Policy									
Guideline Title	ID	Scope and Purpose	Stakeholder Involvement	Rigor of Development	Clarity of Presentation	Applicability	Editorial Independence		Overall Assessment
Critical Issues in the Evaluation and Management of Emergency Department Patients With Suspected Appendicitis	ACEP-14	87%	51%	73%	59%	29%	65%		50%
Critical Issues in the Management of Adult Patients Presenting to the Emergency Department With Community-Acquired Pneumonia	ACEP-15	92%	56%	77%	76%	37%	68%		70%
Neuroimaging and Decisionmaking in Adult Mild Traumatic Brain Injury in the Acute Setting	ACEP-16	93%	58%	78%	82%	44%	72%		83%
Critical Issues in the Evaluation and Management of Adult Patients Presenting to the Emergency Department With Acute Headache	ACEP-17	89%	52%	75%	80%	36%	63%		67%
Critical Issues in the Evaluation and Management of Adult Patients Presenting to the Emergency Department with Acute Heart Failure Syndromes	ACEP-18	91%	52%	75%	77%	27%	58%		73%
Critical Issues in the Evaluation and Management of Adult Patients Presenting to the Emergency Department with Syncope	ACEP-19	90%	53%	76%	73%	25%	37%		70%
Emergency Department Management of Patients Needing Reperfusion Therapy	ACEP-20	91%	60%	82%	81%	43%	68%		80%

Table II. Assessment of the Agree II Six Quality Domains and Overall Assessment of Clinical Policy									
Guideline Title	ID	Scope and Purpose	Stakeholder Involvement	Rigor of Development	Clarity of Presentation	Applicability	Editorial Independence		Overall Assessment
for an ST-Segment Elevation Acute Myocardial Infarction									
Mean ± SD		90 ± 2.8	53 ± 4.2	78 ± 2.7	75 ± 5.2	35 ± 5.7	68 ± 9.2		69 ± 9.1
Coefficient of Variation (CV)		0.03	0.08	0.03	0.07	0.16	0.14		0.13
Range		84 - 96	46 - 63	73 - 82	59 - 82	25 - 46	37 - 77		50 – 83

Table III. Appraiser Recommendation for Clinical Policy Use and Overall Assessment						
Guideline ID	Guideline Title	Five Appraisers				Overall Assessment
		Yes	Yes With Modifications	No		
ACEP-1	Critical Issues in the Diagnosis and Management of the Adult Psychiatric Patient in the Emergency Department	3	2	0		70%
ACEP-2	Critical Issues in the Initial Evaluation and Management of Patients Presenting to the Emergency Department in Early Pregnancy	3	2	0		70%
ACEP-3	Critical Issues in the Evaluation and Management of Adult Patients Presenting to the Emergency Department with Acute Carbon Monoxide Poisoning	3	1	1		63%
ACEP-4	Critical Issues in the Evaluation of Adult Patients with Suspected Transient Ischemic Attack in the Emergency Department	4	1	0		77%
ACEP-5	Clinical Policy for Well-Appearing Infants and Children Younger Than 2 Years of Age Presenting to the Emergency Department With Fever	4	1	0		80%
ACEP-6	Use of Intravenous Tissue Plasminogen Activator for the Management of Acute Ischemic Stroke in the Emergency Department	2	3	0		73%
ACEP-7	Critical Issues in the Evaluation and Management of Adult Patients With Suspected Acute Nontraumatic Thoracic Aortic Dissection	4	1	0		70%
ACEP-8	Critical Issues in the Evaluation and Management of Adult Patients Presenting to the Emergency Department with Seizures	3	2	0		67%
ACEP-9	Procedural Sedation and Analgesia in the Emergency Department	4	1	0		70%
ACEP-10	Critical Issues in the Evaluation and Management of Adult Patients in the Emergency Department with Asymptomatic Elevated Blood Pressure	2	1	2		50%
ACEP-11	Clinical Issues in the Prescribing of Opioids for Adult Patients in the Emergency Department	1	2	2		53%

Table III. Appraiser Recommendation for Clinical Policy Use and Overall Assessment						
Guideline ID	Guideline Title	Five Appraisers				Overall Assessment
		Yes	Yes With Modifications	No		
ACEP-12	Critical Issues in the Evaluation and Management of Adult Patients Presenting to the Emergency Department With Suspected Pulmonary Embolism	4	1	0		73%
ACEP-13	Critical Issues in the Evaluation of Adult Patients Presenting to the Emergency Department With Acute Blunt Abdominal Trauma	4	1	0		70%
ACEP-14	Critical Issues in the Evaluation and Management of Emergency Department Patients With Suspected Appendicitis	1	3	1		50%
ACEP-15	Critical Issues in the Management of Adult Patients Presenting to the Emergency Department With Community-Acquired Pneumonia	4	1	0		70%
ACEP-16	Neuroimaging and Decisionmaking in Adult Mild Traumatic Brain Injury in the Acute Setting	4	1	0		83%
ACEP-17	Critical Issues in the Evaluation and Management of Adult Patients Presenting to the Emergency Department With Acute Headache	2	3	0		67%
ACEP-18	Critical Issues in the Evaluation and Management of Adult Patients Presenting to the Emergency Department with Acute Heart Failure Syndromes	4	1	0		73%
ACEP-19	Critical Issues in the Evaluation and Management of Adult Patients Presenting to the Emergency Department with Syncope	4	1	0		70%
ACEP-20	Emergency Department Management of Patients Needing Reperfusion Therapy for an ST-Segment Elevation Acute Myocardial Infarction	4	1	0		80%
Five appraisers of 20 Guidelines (100 recommendations)		64/100 (64%)	30/100 (30%)	6/100 (6%)		

Table IV. Domain Ratings Before and After ACEP Clinical Policy Methodology Updates			
AGREE II Domain	Pre-Sept 2015 Rating	Post-Sept 2015 Rating	p-value^A
Scope and Purpose	90%	90%	p = 0.67
Stakeholder Involvement	53%	53%	p = 0.94
Rigor of Development	77%	80%	p = 0.02
Clarity of Presentation	74%	77%	p = 0.18
Applicability	33%	38%	p = 0.09
Editorial Independence	66%	73%	p = 0.03
^A Statistical significance was defined as $p < 0.008$ using the Bonferroni correction to account for multiple comparisons			

Table VI(a). Clinical Policy Recommendations Levels, Classes of Evidence, and AGREE II Overall Assessment Rating											
Policy ID	Policy Title	Date of Publication	Number of Recommendations	Recommendation Level			Class of Evidence				Overall Assessment
				Level A	Level B	Level C	Class I	Class II	Class III		
ACEP-20	Emergency Department Management of Patients Needing Reperfusion Therapy for Acute ST-Segment Elevation Myocardial Infarction	Nov 2017	4	0	2	2	0	2	7		80%
ACEP-1	Critical Issues in the Diagnosis and Management of the Adult Psychiatric Patient in the Emergency Department	Apr 2017	4	0	0	4	0	0	6		70%
ACEP-2	Critical Issues in the Initial Evaluation and Management of Patients Presenting to the Emergency Department in Early Pregnancy	Feb 2017	3	0	2	1	0	10	10		63%
ACEP-3	Critical Issues in the Evaluation and Management of Adult Patients Presenting to the Emergency Department with Acute Carbon Monoxide Poisoning	Jan 2017	3	0	3	0	0	6	8		77%
ACEP-4	Critical Issues in the Evaluation of Adult Patients with Suspected Transient Ischemic Attack in the Emergency Department	Sep 2016	4	0	2	2	0	18	45		80%
ACEP-5	Clinical Policy for Well-Appearing Infants and Children Younger Than 2 Years of Age Presenting to the Emergency Department With Fever	May 2016	6	0	2	4	0	3	19		73%

Table VI(a). Clinical Policy Recommendations Levels, Classes of Evidence, and AGREE II Overall Assessment Rating											
Policy ID	Policy Title	Date of Publication	Number of Recommendations	Recommendation Level			Class of Evidence				Overall Assessment
				Level A	Level B	Level C	Class I	Class II	Class III		
ACEP-6	Use of Intravenous Tissue Plasminogen Activator for the Management of Acute Ischemic Stroke in the Emergency Department	Sep 2015	2	0	2	2	1	6	71		73%
ACEP-7	Critical Issues in the Evaluation and Management of Adult Patients With Suspected Acute Nontraumatic Thoracic Aortic Dissection	Jan 2015	6	0	2	4	2	3	23		70%
ACEP-8	Critical Issues in the Evaluation and Management of Adult Patients Presenting to the Emergency Department with Seizures	April 2014	6	1	1	4	1	3	32		67%
ACEP-9	Procedural Sedation and Analgesia in the Emergency Department	Feb 2014	6	1	3	2	2	10	16		70%
ACEP-10	Critical Issues in the Evaluation and Management of Adult Patients in the Emergency Department with Asymptomatic Elevated Blood Pressure	Jul 2013	2	0	0	2	0	2	4		50%
ACEP-11	Clinical Issues in the Prescribing of Opioids for Adult Patients in the Emergency Department	Oct 2012	5	0	1	4	1	4	15		53%
ACEP-12	Critical Issues in the Evaluation and Management of Adult Patients Presenting to the Emergency Department With Suspected Pulmonary Embolism	Jun 2011	10	1	5	4	9	38	75		73%

Table VI(a). Clinical Policy Recommendations Levels, Classes of Evidence, and AGREE II Overall Assessment Rating											
Policy ID	Policy Title	Date of Publication	Number of Recommendations	Recommendation Level			Class of Evidence				Overall Assessment
				Level A	Level B	Level C	Class I	Class II	Class III		
ACEP-13	Critical Issues in the Evaluation of Adult Patients Presenting to the Emergency Department With Acute Blunt Abdominal Trauma	Apr 2011	6	0	3	3	1	7	7		70%
ACEP-14	Critical Issues in the Evaluation and Management of Emergency Department Patients With Suspected Appendicitis	Jan 2010	4	0	3	1	9	21	35		50%
ACEP-15	Critical Issues in the Management of Adult Patients Presenting to the Emergency Department With Community-Acquired Pneumonia	Nov 2009	4	0	2	2	0	10	20		70%
ACEP-16	Neuroimaging and Decisionmaking in Adult Mild Traumatic Brain Injury in the Acute Setting	Dec 2008	5	1	2	2	4	15	24		83%
ACEP-17	Critical Issues in the Evaluation and Management of Adult Patients Presenting to the Emergency Department With Acute Headache	Oct 2008	6	0	3	3	1	10	32		67%
ACEP-18	Critical Issues in the Evaluation and Management of Adult Patients Presenting to the Emergency Department with Acute Heart Failure Syndromes	May 2007	7	0	4	3	5	34	31		73%
ACEP-19	Critical Issues in the Evaluation and Management of Adult Patients Presenting to the Emergency Department with Syncope	Apr 2007	5	2	2	1	4	8	10		70%

Table VI (b). Description of CPG Level Recommendation and Classification of Literature/Evidence ^A			
Level A Recommendation: Generally accepted principles for patient care that reflect a high degree of clinical certainty (e.g., based on evidence from 1 or more Class of Evidence I or multiple Class of Evidence II studies)		Level B Recommendation: Recommendation for patient care that may identify a particular strategy or range of strategies that reflect moderate clinical certainty (e.g., based on evidence from 1 or more Class of Evidence II studies or strong consensus of Class of Evidence III studies)	
		Level C Recommendations: Recommendations for patient care that are based on evidence from Class of Evidence III studies or in the absence of any adequate published literature, based on expert consensus	
Literature/Evidence Classification Schema			
Design Class	Therapy	Diagnosis	Prognosis
I	Randomized, controlled trial or meta-analysis of randomized trials	Prospective cohort using a criterion standard or meta-analysis of prospective studies	Population prospective cohort or meta-analysis of prospective studies
II	Nonrandomized trial	Retrospective observational	Retrospective cohort. Case control
III	Case series	Case series	Case series
^A The recommendation levels and Design Class (I, II,III) are from ACEP Clinical Policy Development Process https://www.acep.org/Clinical---Practice-Management/ACEP-Clinical-Policies-Development-Process/#sm.00000nl048nlvdoawtf22kfngb0xo The Design Class descriptions are found in Appendix A of the Clinical Policies			

Appendix

Appendix I: The AGREE II Instrument

Domain 1. Scope and Purpose is concerned with the overall aim of the guideline, the specific health questions and the target population (items 1-3).

- 1) Objectives(s) – The overall objectives and guidelines is (are) specifically described
- 2) Question(s) – The health question(s) is (are) specifically described
- 3) Population – (patients, public, etc) to whom the guideline is meant to apply is specifically described

Domain 2. Stakeholder Involvement focuses on the extent to which the guideline was developed by the appropriate stakeholders and represents the views of its intended users (items 4 -6).

- 4) Guidelines Group – the guideline development group includes individuals from all relevant professional groups
- 5) Patient Preferences – The view / preferences of the target population (patients, public, etc) have been sought
- 6) Target – The target users of the guideline are clearly defined

Domain 3. Rigor of Development relates to the process used to gather and synthesize the evidence and the methods to formulate and update the recommendations (items 7-14).

- 7) Systematic Methods – Systematic methods were used to search for evidence
- 8) Selection Criteria – The criteria for selecting the evidence are clearly described
- 9) Strength and Limitations – The strengths and limitations of the body of evidence are clearly described
- 10) Method of Recommendations – The methods for formulating the recommendations are clearly described
- 11) Benefits, Side Effects and Risks – These aspects have been considered in formulating the recommendations
- 12) Evidence Link – There is an explicit link between the recommendations and the supporting evidence
- 13) External Review – The guideline has been externally reviewed by experts prior to its publication
- 14) Update Procedures – A procedure for updating the guideline is provided

Domain 4. Clarity of Presentation deals with the language, structure, and format of the guideline (items 15-17).

- 15) Recommendation Specific - The recommendations are specific and unambiguous
- 16) Options for Management – Different options for management of the condition/health issue are clearly presented
- 17) Recommendations identifiable – Key recommendation are easily identifiable

Domain 5. Applicability pertains to the likely barriers and facilitators to implementation, strategies to improve uptake, and resource implications of applying the guideline (items 18-21).

- 18) Facilitators and Barriers – The guideline describes facilitators and barriers to its application
- 19) Tools – The guideline provides advice and/or tools on how the recommendation can be put into practice
- 20) Resource Implications – Potential resource implications of applying the recommendation have been considered
- 21) Monitoring/Audit Criteria – The guideline presents monitoring and/or auditing criteria.

Domain 6. Editorial Independence is concerned with the formulation of recommendations not being unduly biased with competing interests (items 22-23).

- 22) Funding Body – The views of the funding body have not influenced the content of the guideline
- 23) Competing Interests – Competing interests of guideline group members have been recorded and addressed.

Overall Assessment includes the rating of the

- The overall quality of the guideline (1-7 scale)
- Whether the guideline would be recommended for use in practice (yes, yes with modification, or no)

Appendix II

AGREE II Domains and Overall Assessment Ratings by AZ								
Clinical Policy	ID	Scope and Purpose (21)^A	Stake Holder Involvement (21)	Rigor of Development (56)	Clarity of Presentation (21)	Applicability (28)	Editorial Independence (14)	Overall Quality (7)
Critical Issues in the Diagnosis and Management of the Adult Psychiatric Patient in the Emergency Department	1	20 (94%) ^B	14 (61%)	48 (83%)	17 (78%)	10 (25%)	10 (67%)	5 (67%)
Critical Issues in the Initial Evaluation and Management of Patients Presenting to the Emergency Department in Early Pregnancy	2	20 (94%)	14 (61%)	48 (83%)	17 (78%)	7 (13%)	10 (67%)	5 (67%)
Critical Issues in the Evaluation and Management of Adult Patients Presenting to the Emergency Department with Acute Carbon Monoxide Poisoning	3	20 (94%)	13 (56%)	48 (83%)	17 (78%)	7 (13%)	10 (67%)	5 (67%)
Critical Issues in the Evaluation of Adult Patients with Suspected Transient Ischemic Attack in the Emergency Department	4	20 (94%)	12 (50%)	48 (83%)	17 (78%)	9 (21%)	10 (67%)	5 (67%)
Clinical Policy for Well-Appearing Infants and Children Younger Than 2 Years of Age Presenting to the Emergency Department With Fever	5	20 (94%)	14 (61%)	48 (83%)	17 (78%)	8 (17%)	10 (67%)	5 (67%)
Use of Intravenous Tissue Plasminogen Activator for the Management of Acute Ischemic Stroke in the Emergency Department	6	20 (94%)	11 (44%)	45 (77%)	17 (78%)	9 (21%)	10 (67%)	5 (67%)
Critical Issues in the Evaluation and Management of Adult Patients With Suspected Acute Nontraumatic Thoracic Aortic Dissection	7	20 (94%)	13 (56%)	48 (83%)	17 (78%)	7 (13%)	10 (67%)	5 (67%)
Critical Issues in the Evaluation and	8	20	14	46	17	8	10	5

AGREE II Domains and Overall Assessment Ratings by AZ								
Clinical Policy	ID	Scope and Purpose (21)^A	Stake Holder Involvement (21)	Rigor of Development (56)	Clarity of Presentation (21)	Applicability (28)	Editorial Independence (14)	Overall Quality (7)
Management of Adult Patients Presenting to the Emergency Department with Seizures		(94%)	(61%)	(83%)	(78%)	(17%)	(67%)	(67%)
Procedural Sedation and Analgesia in the Emergency Department	9	20 (94%)	14 (61%)	46 (79%)	17 (78%)	8 (17%)	10 (67%)	5 (67%)
Critical Issues in the Evaluation and Management of Adult Patients in the Emergency Department with Asymptomatic Elevated Blood Pressure	10	20 (94%)	13 (56%)	45 (77%)	17 (78%)	8 (17%)	10 (67%)	5 (67%)
Clinical Issues in the Prescribing of Opioids for Adult Patients in the Emergency Department	11	20 (94%)	19 (89%)	46 (79%)	17 (78%)	8 (17%)	12 (83%)	5 (67%)
Critical Issues in the Evaluation and Management of Adult Patients Presenting to the Emergency Department With Suspected Pulmonary Embolism	12	20 (94%)	13 (56%)	46 (79%)	17 (78%)	6 (8%)	10 (67%)	5 (67%)
Critical Issues in the Evaluation of Adult Patients Presenting to the Emergency Department With Acute Blunt Abdominal Trauma	13	20 (94%)	13 (56%)	46 (79%)	17 (78%)	6 (8%)	10 (67%)	5 (67%)
Critical Issues in the Evaluation and Management of Emergency Department Patients With Suspected Appendicitis	14	20 (94%)	13 (56%)	46 (79%)	17 (78%)	7 (13%)	8 (50%)	5 (67%)
Critical Issues in the Management of Adult Patients Presenting to the Emergency Department With Community-Acquired Pneumonia	15	20 (94%)	15 (67%)	46 (79%)	17 (78%)	7 (13%)	8 (50%)	5 (67%)
Neuroimaging and Decision making in Adult Mild Traumatic Brain Injury in the Acute Setting	16	20 (94%)	13 (56%)	44 (75%)	17 (78%)	7 (13%)	12 (83%)	5 (67%)

